

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 10:44:58 ; Search time 7578 Seconds

(without alignments)
12239.632 Million cell updates/sec

Title: US-09-900-751-1

Sequence: 1 catgctagcagcgtccgcgg.....ttaaaaaaaaaaaaaaaaaa 3106

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rdg:*
26: em_gss_pbg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3102.8	99.9	3248	11	BC008514 Mus muscu
2	3087.6	99.4	4017	11	AK052738 Mus muscu
3	1790	57.6	2348	29	AY419858 Mus muscu
4	1558.6	50.2	2487	29	AY419856 Homo sapi

5	958.2	30.8	2487	29	AY419857 Pan trogl
6	956.4	30.8	2689	11	AK085410 Mus muscu
7	843.8	27.2	944	13	BQ885551 AGENCOURT
8	833.6	26.8	884	12	B1688367
9	788.8	25.4	803	13	B1690615
10	778.2	25.1	930	13	BQ951529 AGENCOURT
11	760.4	24.5	762	12	BG919617
12	759.4	24.4	874	14	CB208872
13	747.2	24.1	798	14	B1664081
14	739.2	23.8	746	13	B1756896
15	739.2	23.8	846	12	B1416218
16	735.8	23.7	1116	10	BF301347
17	732	23.6	786	12	B1646380
18	721	23.2	928	12	BG915045
19	720.8	23.2	777	12	B1150180
20	716	23.1	928	12	B1694441
21	710.8	22.9	838	12	B1656113
22	707.4	22.8	767	12	B1695744
23	706.8	22.8	840	12	B1453178
24	706.2	22.7	939	14	CA976536
25	692.2	22.3	725	12	B1248406
26	689.6	22.2	960	12	B1414245
27	687.2	22.1	833	10	BF540669
28	684.4	22.0	697	12	B1853355
29	681.8	22.0	708	12	B1078327
30	681.2	21.9	838	13	B1078327
31	678.2	21.8	1035	10	B1119461
32	671	21.6	671	12	BQ861619
33	665.4	21.4	836	13	B0517436
34	658.6	21.2	728	12	B1112917
35	651	21.0	651	14	CF169769
36	650.4	20.9	927	10	BF578780
37	647.4	20.8	649	12	B1697395
38	643.8	20.7	698	13	BY756283
39	633.6	20.4	886	12	B1657843
40	630.4	20.3	744	12	B1454477
41	623.8	20.1	928	10	BF584458
42	623.4	20.1	854	12	B1416184
43	621.4	20.0	765	10	B116535
44	617.2	19.9	630	10	BF320582
45	613	19.7	613	14	CA551053

ALIGNMENTS

RESULT 1
LOCUS BC008514 3248 bp mRNA linear HTC B17-DEC-2003
DEFINITION Mus musculus suppression of tumorigenicity 14 (colon carcinoma),
ACCESSION BC008514
VERSION BC008514.1 GI:14250187
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Dergs,J.G.,
Klausner,R.D., Collins,F.S., Wagner,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Udell,T.B., Toshyk,L.S.,
Carninci,P., Prange,C., Raha,S.S., Loguclano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McKean,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Bulky,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Riley,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

TITLE
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
12477932
2 (bases 1 to 3248)
Strausberg, R.
Direct Submission
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Lotmar Hemilghausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mdcpaxil.stanford.edu
Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 7 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6677898
This clone has the following problem: no 5' EST match.

FEATURES
source
location/qualifiers

1..3248
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3489914"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old. gross tissue."
/clone_id="NCI CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 99.9%; Score 3102.8; DB 11; Length 3248;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATGGTAGACGGCTGCCCCGAGAGGACACGCGTGTGAGACCGGGGATGACCCGCCAAA 60
DB 111 CATGGTAGACGGCTGCCCCGAGAGGACACGCGTGTGAGACCGGGGATGACCCGCCAAA 170
QY 61 CCAATGGTAGCAATGGGGGCGGCAAGCGCGAGGGGGCTTCAGGACTTCGGCGGGGAGC 120
DB 171 CCAATGGTAGCAATGGGGGCGGCAAGCGCGAGGGGGCTTCAGGACTTCGGCGGGGAGC 230
QY 121 TCAAGTACCACTCCCGGCTAGAGAACATGAATGGCTTTAGAGAGGGGTGTGAGTTCTGTC 180
DB 221 TCAAGTACCACTCCCGGCTAGAGAACATGAATGGCTTTAGAGAGGGGTGTGAGTTCTGTC 290
QY 181 CTGGCAACATGCGCAAGAAAGTGAGAGCGAGGCCCGGCGCTGGGTGTGTCTGTGG 240
DB 291 CTGGCAACATGCGCAAGAAAGTGAGAGCGAGGCCCGGCGCTGGGTGTGTCTGTGG 350
QY 241 CAGTGTCTTGAAGCTTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 300
DB 351 CAGTGTCTTGAAGCTTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 410
QY 301 ATTATCGAATGTGGGGTTCAAAAAGTCTTCATATGCGCATCTGAGGATCACAATGAGA 360

DB 411 ATTATCGAATGTGGGGTTCAAAAAGTCTTCATATGCGCATCTGAGGATCACAATGAGA 470
QY 361 TCTTTCTGATGCGATATGAGAACTCACTCCACAGATTATACAGCTGGCCAGCAG 420
DB 471 TCTTTCTGATGCGATATGAGAACTCACTCCACAGATTATACAGCTGGCCAGCAG 530
QY 421 TGAAGAGGCGCTGAGAGCTGTGTACATGAATGCTCTGTGGTCCCTACCAAGA 480
DB 531 TGAAGAGGCGCTGAGAGCTGTGTACATGAATGCTCTGTGGTCCCTACCAAGA 590
QY 481 AGTGGCTGTAACTGGCTTCACTGAGAGGAGAGTCAATGCGCTACTGTGTGAGTTCA 540
DB 591 AGTGGCTGTAACTGGCTTCACTGAGAGGAGAGTCAATGCGCTACTGTGTGAGTTCA 650
QY 541 GCATCCCCCAACCTGGGAGAGAGAGTGTATGAGCCCATAGGCTGTGAGCGAGTTGAA 600
DB 651 GCATCCCCCAACCTGGGAGAGAGAGTGTATGAGCCCATAGGCTGTGAGCGAGTTGAA 710
QY 601 CATTGCCACCCCGAGACAGGGCACTGAATCTCTTGTGTAACTGTGTGAGCTTCC 660
DB 711 CATTGCCACCCCGAGACAGGGCACTGAATCTCTTGTGTAACTGTGTGAGCTTCC 770
QY 661 CCAATTGACCCCAATATGCTGCAAGAGACTGAGCAACAAGTGTGAGTTGGCTGATG 720
DB 771 CCAATTGACCCCAATATGCTGCAAGAGACTGAGCAACAAGTGTGAGTTGGCTGATG 830
QY 721 CCCATGGTGAGAGAGTACAGCTTCACTACCCCTGCTTCCCAACAGTCTCTACCCG 780
DB 831 CCCATGGTGAGAGAGTACAGCTTCACTACCCCTGCTTCCCAACAGTCTCTACCCG 890
QY 781 CGCATGCGCGCTGCGCATGCTGCTGCGGGGAGACGCTCACTGTGTGAGCTTCACT 840
DB 891 CGCATGCGCGCTGCGCATGCTGCTGCGGGGAGACGCTCACTGTGTGAGCTTCACT 950
QY 841 TCCGAACCTTGAATGTCGCTCCCTGATAGAGCATGCGAGTCACTGTGTGAGTATG 900
DB 951 TCCGAACCTTGAATGTCGCTCCCTGATAGAGCATGCGAGTCACTGTGTGAGTATG 1010
QY 901 ATAGCCTGAGCCCGCATGAGAACCCCAAGCTGTGTGAGCTGTGTGAGCACTTCTACCT 960
DB 1011 ATAGCCTGAGCCCGCATGAGAACCCCAAGCTGTGTGAGCTGTGTGAGCACTTCTACCT 1070
QY 961 CTTAACAACCTGACTTCTCTCTCTCCCAAGAACGCTTCTTGTGTGAGCTGTGTGAG 1020
DB 1071 CTTAACAACCTGACTTCTCTCTCTCCCAAGAACGCTTCTTGTGTGAGCTGTGTGAG 1130
QY 1021 CTGACCGGCGCATCTGCTGCTCTCTCTCCCAAGAACGCTTCTTGTGTGAGCTGTGTGAG 1080
DB 1131 CTGACCGGCGCATCTGCTGCTCTCTCTCCCAAGAACGCTTCTTGTGTGAGCTGTGTGAG 1190
QY 1081 GTGCGGCTTTTGTAGTGACACCCCAAGGACATTTAGACGCTTCTTGTGTGAGCTGTGTGAG 1140
DB 1191 GTGCGGCTTTTGTAGTGACACCCCAAGGACATTTAGACGCTTCTTGTGTGAGCTGTGTGAG 1250
QY 1141 ACCGCGCCCAATCACTGCACTGCAATGAAATATCAAGGTGCCCAACACCGGAACGTAAG 1200
DB 1251 ACCGCGCCCAATCACTGCACTGCAATGAAATATCAAGGTGCCCAACACCGGAACGTAAG 1310
QY 1201 TGGCGTTCAACCTTTCTTATCTGTGTGAGACCCCAAGTACCAAGTGGGCTCTGCAACAA 1260
DB 1311 TGGCGTTCAACCTTTCTTATCTGTGTGAGACCCCAAGTACCAAGTGGGCTCTGCAACAA 1370
QY 1261 ACTATGTGAGATCAACCGGGAGAAAGTACTGCGGTGAGAGGTCCAGTTTGTGTGAGCA 1320
DB 1371 ACTATGTGAGATCAACCGGGAGAAAGTACTGCGGTGAGAGGTCCAGTTTGTGTGAGCA 1430
QY 1321 GCAACAGACGACAGATTACAGTTCACCTTCTGATTCATTCGTAACCGGACACCGGCT 1380
DB 1431 GCAACAGACGACAGATTACAGTTCACCTTCTGATTCATTCGTAACCGGACACCGGCT 1490
QY 1381 TCCATGTGAGATCACTTCTCTCAAGCTCCACGACCCGCTCCAGGATGTTCAATGTGA 1440

Db 1491 TCCTAGCTGAGTACTCTCTCTACAGCTCAACGACCCGCTGCCAGGATGTTTCACTGCA 1550
 QY 1441 AGACTGACGGTGCATCCGAAAGAACTGGCTGCGACGCTGGGCACTGCCGAT 1500
 Db 1551 AGACTGACGGTGCATCCGAAAGAACTGGCTGCGACGCTGGGCACTGCCGAT 1610
 QY 1501 ATAGTATGACGGTGCATCCGAAAGAACTGGCTGCGACGCTGGGCACTGCCGAT 1560
 Db 1611 ATAGTATGACGGTGCATCCGAAAGAACTGGCTGCGACGCTGGGCACTGCCGAT 1670
 QY 1561 TCTGCAAGCCCTCTTCTGGGCTCTGTCAGTGTCAAGTGTCAAGTGTGGGAGCGAAGTGA 1620
 Db 1671 TCTGCAAGCCCTCTTCTGGGCTCTGTCAGTGTCAAGTGTGGGAGCGAAGTGA 1730
 QY 1621 AGAGGGCTGACGCTGTCTCTGGGAGTTTCAAGTGTTCATGGGAAGTGTCTCCCTC 1680
 Db 1731 AGAGGGCTGACGCTGTCTCTGGGAGTTTCAAGTGTTCATGGGAAGTGTCTCCCTC 1790
 QY 1681 AGAGCCAGAGTGTATGGGAGAGCACTGTGGAGATGGGTCTGACGAGGCTTCATGTG 1740
 Db 1791 AGAGCCAGAGTGTATGGGAGAGCACTGTGGAGATGGGTCTGACGAGGCTTCATGTG 1850
 QY 1741 ACAGCGTGAATGTCTCTCTTGGACCAATATATCTACCGCTGCCAAATGGCTCTGTC 1800
 Db 1851 ACAGCGTGAATGTCTCTCTTGGACCAATATATCTACCGCTGCCAAATGGCTCTGTC 1910
 QY 1801 TGAAGCAGGCAACCTGTAGTGTATGGAAGAGCGAGCTGAGCGATGGCTCGATGAGA 1860
 Db 1911 TGAAGCAGGCAACCTGTAGTGTATGGAAGAGCGAGCTGAGCGATGGCTCGATGAGA 1970
 QY 1861 AAAAAGTGAATGTGGGCTGCGATCCTTTACCAAAAGAGCTGGCGTGTGGTGGAGCA 1920
 Db 1971 AAAAAGTGAATGTGGGCTGCGATCCTTTACCAAAAGAGCTGGCGTGTGGTGGAGCA 2030
 QY 1921 ATGCGAGCAGGAGCGAGTGGCTCTGGCAGAGTGAAGCTCCAGCCCTGGGCAAGGCACT 1980
 Db 2031 ATGCGAGCAGGAGCGAGTGGCTCTGGCAGAGTGAAGCTCCAGCCCTGGGCAAGGCACT 2090
 QY 1981 TGTGTGGGCTCTGGCTCATCTCTCTGACGTGGTCTCTGACGTCAATGCTTTCAAG 2040
 Db 2091 TGTGTGGGCTCTGGCTCATCTCTCTGACGTGGTCTCTGACGTCAATGCTTTCAAG 2150
 QY 2041 ATGACAAAATTTCAAGTACTCAGACTACAGATGAGAGCGGCTTCTGGGCTGTGG 2100
 Db 2151 ATGACAAAATTTCAAGTACTCAGACTACAGATGAGAGCGGCTTCTGGGCTGTGG 2210
 QY 2101 ACCAGAGCAGGAGCGAGTGGCTCTGGGAGTGCAGAGTGAAGTCAAAAGTATCATCAACC 2160
 Db 2211 ACCAGAGCAGGAGCGAGTGGCTCTGGGAGTGCAGAGTGAAGTCAAAAGTATCATCAACC 2270
 QY 2161 ACCCTTCTTCAATGATTTTCACTTTCGATGATGACATGCGCTTGTGAGCTGAGAAAT 2220
 Db 2271 ACCCTTCTTCAATGATTTTCACTTTCGATGATGACATGCGCTTGTGAGCTGAGAAAT 2330
 QY 2221 CGGTGAGTGAAGCAGCGTGGGAGCGGCTTGTGCTGCTGATGCTTCAATGCTTCTTC 2280
 Db 2331 CGGTGAGTGAAGCAGCGTGGGAGCGGCTTGTGCTGCTGATGCTTCAATGCTTCTTC 2390
 QY 2281 CTGCTGGCAGGCGCATCTGGGTCAAGAGCTGGGAGCACAAGAGAGGAGTACCGAG 2340
 Db 2391 CTGCTGGCAGGCGCATCTGGGTCAAGAGCTGGGAGCACAAGAGAGGAGTACCGAG 2450
 QY 2341 CGCTGATCTTGCAGAAAGGCTGATGCTGCTGATCATCAACCAACCACTGTGAGAGCTCA 2400
 Db 2451 CGCTGATCTTGCAGAAAGGCTGATGCTGCTGATCATCAACCAACCACTGTGAGAGCTCA 2510
 QY 2401 TGCCGACAGCAGATCAACCCCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 Db 2511 TGCCGACAGCAGATCAACCCCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2570
 QY 2461 CTGCGCAGGCTGACTCTGTGGTGGCTTGTCAAGCGCGAGAAAGATGGGCAATGTTCC 2520
 Db 2571 CTGCGCAGGCTGACTCTGTGGTGGCTTGTCAAGCGCGAGAAAGATGGGCAATGTTCC 2630

QY 2521 AGCTGATGTGATGAGCTGGGAGTGAAGGCTGCCTCAGAGAGAAACAAGCCAGGCTGTACA 2580
 Db 2631 AGCTGATGTGATGAGCTGGGAGTGAAGGCTGCCTCAGAGAGAAACAAGCCAGGCTGTACA 2590
 QY 2581 CAAGGCTCCTGTAGTGTGGGACTGAGATCAAGAGACACTGGGATATAGCAGATGAC 2640
 Db 2691 CAAGGCTCCTGTAGTGTGGGACTGAGATCAAGAGACACTGGGATATAGCAGATGAC 2750
 QY 2641 AGACAGCCCAACCAACCAACCAAGAGATGCCGACATGACACCTGATACAGAGAG 2700
 Db 2751 AGACAGCCCAACCAACCAACCAAGAGATGCCGACATGACACCTGATACAGAGAG 2810
 QY 2701 GAACCTGACGACATTTATGCTGGCTCCCGCCCGCAACAACCAACCACTGTGAACT 2760
 Db 2811 GAACCTGACGACATTTATGCTGGCTCCCGCCCGCAACAACCAACCACTGTGAACT 2870
 QY 2761 GCATCCTTAGACTCAGAGTTCCTTCCAAAGTGGGAGCCCTCAAGAGTTGAGAGAGAC 2820
 Db 2871 GCATCCTTAGACTCAGAGTTCCTTCCAAAGTGGGAGCCCTCAAGAGTTGAGAGAGAC 2930
 QY 2821 TTGCGCTAGAGGCGCCACCTGGGAGCAAGGTTTGAATGAGAGCTTCCCTTACGCC 2880
 Db 2931 TTGCGCTAGAGGCGCCACCTGGGAGCAAGGTTTGAATGAGAGCTTCCCTTACGCC 2990
 QY 2881 CTGAGCTGGTGAAGATGATGCTGCTCCGAGAGGCTTCCAACTGATGAGTCC 2940
 Db 2991 CTGAGCTGGTGAAGATGATGCTGCTCCGAGAGGCTTCCAACTGATGAGTCC 3050
 QY 2941 GGGAGCCTTATGGAGAGAGGCTCAAGGCTCACTTTTACAGAAAGCGCAGCCTTAGA 3000
 Db 3051 GGGAGCCTTATGGAGAGAGGCTCAAGGCTCACTTTTACAGAAAGCGCAGCCTTAGA 3110
 QY 3001 ACCCAGAAAAGAGTGAACCTTAAGCTGAATGTTTGTGCTGCTCCAGGCTGTAT 3060
 Db 3111 ACCCAGAAAAGAGTGAACCTTAAGCTGAATGTTTGTGCTGCTCCAGGCTGTAT 3170
 QY 3061 TTGAGAGTAAACATTTATTTCTTTTAAAAA 3106
 Db 3171 TTGAGAGTAAACATTTATTTCTTTAAAAA 3216

RESULT 2
 AK052738
 LOCUS
 DEFINITION
 Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched library, clone: D63041B06 product: suppression of tumorigenicity 14 (colon carcinoma), full insert sequence.
 ACCESSION
 AK052738
 VERSION
 AK052738.1 GI:26342936
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 Carninci, P., and Hayashizaki, Y.
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 PUBMED
 REFERENCE
 2
 10349636
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 3
 11042159
 AUTHORS
 Shihara, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, T., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

QY 841 TCGAAGCTTTGATGTCCTCCCTGTGATGAGCATGGCAGTGAACCTGGTCAACCGTGTATG 900
| | | | |
Db 952 TCGGAACCTTTGATGTGTGCTCCCTGTATGATGAGCATGGCAGTGAACCTGGTCAACCGTGTATG 1011
| | | | |
QY 901 ATAGCTGAGCCCCATGAAACCCCAAGCTGTGTGAGCGGCTGTGTGACCTTCTCAACCT 960
| | | | |
Db 1012 ATAGCTGAGCCCCATGAAACCCCAAGCTGTGTGAGCGGCTGTGTGACCTTCTCAACCT 1071
| | | | |
QY 961 CCTACAACTGACTTTCCTCTCTCCCAAGACGTCTTCTGTGTACAGCTGATTAACCAATA 1020
| | | | |
Db 1072 CCTACAACTGACTTTCCTCTCTCCCAAGACGTCTTCTGTGTACAGCTGATTAACCAATA 1131
| | | | |
QY 1021 CTGACCGGAGCATCTCCGCTTGTGAGGCACTTCTTCCAGCTGCCCAAGATGAGCACT 1080
| | | | |
Db 1132 CTGACCGGAGCATCTCCGCTTGTGAGGCACTTCTTCCAGCTGCCCAAGATGAGCACT 1191
| | | | |
QY 1081 GTGGCGGCTTTTGTAGTGACACCCCAAGGACATTTTGTAGAGCCCTCTCTATCCAGGCACT 1140
| | | | |
Db 1192 GTGGCGGCTTTTGTAGTGACACCCCAAGGACATTTTGTAGAGCCCTCTCTATCCAGGCACT 1251
| | | | |
QY 1141 ACCCGCCCAACATCACTGACATGGAATATCAAGGTGCCCAACACCGGAAAGTGAAG 1200
| | | | |
Db 1252 ACCCGCCCAACATCACTGACATGGAATATCAAGGTGCCCAACACCGGAAAGTGAAG 1311
| | | | |
QY 1201 TGGGCTTCAAACTCTTCTATCTGTGTGAGCCCAAGTACAGGGGCTCTGTACCAAG 1260
| | | | |
Db 1312 TGGGCTTCAAACTCTTCTATCTGTGTGAGCCCAAGTACAGGGGCTCTGTACCAAG 1371
| | | | |
QY 1261 ACTATGTGAGATCAACGGGAGAAAGTACTGGGTGTAGAGTCCAGTTTGTGTAGCA 1320
| | | | |
Db 1372 ACTATGTGAGATCAACGGGAGAAAGTACTGGGTGTAGAGTCCAGTTTGTGTAGCA 1431
| | | | |
QY 1321 GGAACAGCAGCAAGATTACATTCACCTTCTATCTGTATCACTGTACACGACACCGGAT 1380
| | | | |
Db 1432 GGAACAGCAGCAAGATTACATTCACCTTCTATCTGTATCACTGTATCAACGACACCGGAT 1491
| | | | |
QY 1381 TCTTACTGAGTACTCTCTCTACGATCCCAAGACCCCGCCCAAGGATTTCAATGTGA 1440
| | | | |
Db 1492 TCTTACTGAGTACTCTCTCTACGATCCCAAGACCCCGCCCAAGGATTTCAATGTGA 1551
| | | | |
QY 1441 AGACTGAGCGGTGATCCGAAAGAAAGTACGCTGTGAGCGGCTGTGAGCACTGCCGAT 1500
| | | | |
Db 1552 AGACTGAGCGGTGATCCGAAAGAAAGTACGCTGTGAGCGGCTGTGAGCACTGCCGAT 1611
| | | | |
QY 1501 ATAGTATGAGGCTTACTGCGGATGCAATGCAACCCACCACTTCACTGTCAAAAAACA 1560
| | | | |
Db 1612 ATAGTATGAGGCTTACTGCGGATGCAATGCAACCCACCACTTCACTGTCAAAAAACA 1671
| | | | |
QY 1561 TCTGCAAGCCCTCTTCTGGGCTGTGTGACAGTGTCAAGCTGTGAGCACTGTGAGCA 1620
| | | | |
Db 1672 TCTGCAAGCCCTCTTCTGGGCTGTGTGACAGTGTCAAGCTGTGAGCACTGTGAGCA 1731
| | | | |
QY 1621 AGGAGGCTGACGCTGTCTGTGTGAGTTTCAATGTTCATGTGAAAGTGTCTCCCTC 1680
| | | | |
Db 1732 AGGAGGCTGACGCTGTCTGTGTGAGTTTCAATGTTCATGTGAAAGTGTCTCCCTC 1791
| | | | |
QY 1681 AGAGCCAGAAATGTATGTGAAAGGACAACTGTGAGATGAGTCTGTGACGAGGCTTCA 1740
| | | | |
Db 1792 AGAGCCAGAAATGTATGTGAAAGGACAACTGTGAGATGAGTCTGTGACGAGGCTTCA 1851
| | | | |
QY 1741 ACAGCGTGAATGTGTCTCTTGTGACCAAAATATACCTTCCGTCGCAAAATGCTCTGT 1800
| | | | |
Db 1852 ACAGCGTGAATGTGTCTCTTGTGACCAAAATATACCTTCCGTCGCAAAATGCTCTGT 1911
| | | | |
QY 1801 TGAGCAAGGAGCAACCTGAGTGTATGTGAGAGAGAGCACTGTAGGATGAGTCCGATGAG 1860
| | | | |
Db 1912 TGAGCAAGGAGCAACCTGAGTGTATGTGAGAGAGAGCACTGTAGGATGAGTCCGATGAG 1971
| | | | |
QY 1861 AAAAATGTGACTGTGGCTGTGCAATCTTTTACAAAACAGGCTGTGCTGTGTGTGTGAG 1920
| | | | |
Db 1972 AAAAATGTGACTGTGGCTGTGCAATCTTTTACAAAACAGGCTGTGCTGTGTGTGTGAG 2031
| | | | |
QY 1921 ATGCGAGAGAGGCTGAGTGTGCTGTGAGTGTGAGCTTCAAGCCCTGTGGCCAGGCACT 1980
| | | | |

Db 2032 ATGCGAGAGAGGCTGAGTGTGCTCCCTGAGAGTGAAGCTTCCAGGCTGTGGGCAAGGCACT 2091
| | | | |
QY 1981 TGTGTGGGCTGTGCTATCTCTCTGTGATGCTGTGTCTCTGTGAGTCAATGTCTTCAAG 2040
| | | | |
Db 2092 TGTGTGGGCTGTGCTATCTCTCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2151
| | | | |
QY 2041 ATGACAAAAATTTCAAGTACTGAGATCAAGATGTGAGAGGCTTCTGTGGTCTGTGTG 2100
| | | | |
Db 2152 ATGACAAAAATTTCAAGTACTGAGATCAAGATGTGAGAGGCTTCTGTGGTCTGTGTG 2211
| | | | |
QY 2101 ACCAGAGCAAGGCACTGTCTGTGGGTGTGAGAGTGAAGTCAACATGATATCAACCC 2160
| | | | |
Db 2212 ACCAGAGCAAGGCACTGTCTGTGGGTGTGAGAGTGAAGTCAACATGATATCAACCC 2271
| | | | |
QY 2161 ACCCTTCTTCAATGATTTTCACTTGTGATATGATATGCTGTGTGTGTGTGTGTGTGT 2220
| | | | |
Db 2272 ACCCTTCTTCAATGATTTTCACTTGTGATATGATATGCTGTGTGTGTGTGTGTGTGT 2331
| | | | |
QY 2221 CGGTGAGTACAGACCGTGTGTGCGCCCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 2280
| | | | |
Db 2332 CGGTGAGTACAGACCGTGTGTGCGCCCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 2391
| | | | |
QY 2281 CTGTGTGAGAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
| | | | |
Db 2392 CTGTGTGAGAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2451
| | | | |
QY 2341 CGGTGATCTGTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
| | | | |
Db 2452 CGGTGATCTGTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2511
| | | | |
QY 2401 TGCAGAGAGATCAACCCCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
| | | | |
Db 2512 TGCAGAGAGATCAACCCCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2571
| | | | |
QY 2461 CTGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520
| | | | |
Db 2572 CTGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2631
| | | | |
QY 2521 AGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
| | | | |
Db 2632 AGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2691
| | | | |
QY 2581 CAAGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
| | | | |
Db 2692 CAAGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2751
| | | | |
QY 2641 AGACAGCCGACCAACCAACCAAGGATGCCGACATGTGACACCTGTGATACAGAGAG 2700
| | | | |
Db 2752 AGACAGCCGACCAACCAACCAAGGATGCCGACATGTGACACCTGTGATACAGAGAG 2811
| | | | |
QY 2701 GAAACCTGACGCAATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
| | | | |
Db 2812 GAAACCTGACGCAATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2871
| | | | |
QY 2761 GCATCTTGTAGACTCAGAGTCTTTCAAAATGTGAGAGGCTTCAAGAGTGTGAGAGAGAC 2820
| | | | |
Db 2872 GCATCTTGTAGACTCAGAGTCTTTCAAAATGTGAGAGGCTTCAAGAGTGTGAGAGAGAC 2931
| | | | |
QY 2821 TTGCTGTGTAGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
| | | | |
Db 2932 TTGCTGTGTAGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2991
| | | | |
QY 2881 CTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
| | | | |
Db 2992 CTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3051
| | | | |
QY 2941 GGGAGCCCTATGTGAGAGAGGAGGCTCAGAGTCACTTTTTCAGAAAGGCGGCAAGCCCTGAG 3000
| | | | |
Db 3052 GGGAGCCCTATGTGAGAGAGGAGGCTCAGAGTCACTTTTTCAGAAAGGCGGCAAGCCCTGAG 3111
| | | | |
QY 3001 ACCCCAGAAAAGAGTGTATCTTAAAGCTGAAATTTTGTGTGTGTGTGTGTGTGTGTGTGT 3060
| | | | |

Db 3112 ACCCGAAGAAAGAGTGTACCTTAAGCTCAATGTGTTGCTGTTCGCCAGGGGTGGGTAT 3171
QY 3061 TTGAGAGTAAACATTTTATTTCTTTTAAAAA 3094
Db 3172 TTGAGAGTAAACATTTTATTTCTTTTGGTAA 3205

RESULT 3
AY419858 2348 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus ST14 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY419858
VERSION AY419858.1 GI:39775815
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2348)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Perriere, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL PUBMED 14671302
TITLE 2 (bases 1 to 2348)
REFERENCE Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Perriere, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment
FEATURES
Location/Qualifiers
source 1..2348
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..2348
/gene="ST14"
/locus_tag="HCM7030"

ORIGIN
Query Match 57.6%; Score 1790; DB 29; Length 2348;
Best Local Similarity 78.0%; Pred. No. 0;
Matches 1939; Conservative 0; Mismatches 409; Indels 139; Gaps 1;
gene

QY 144 AACATGATGCTTTGAGAGGCTGTGAGTTCTGCTGCGAACAATGCCAAGAAAGT 203
Db 1 AACATGATGCTTTGAGAGGCTGTGAGTTCTGCTGCGAACAATGCCAAGAAAGT 60
QY 204 GAGAAGGAGGCCCCAGGCGCTGGGTGTGCTGTGGAGAGTGTCTTCTCTG 263
Db 61 GAGAAGGAGGCCCCAGGCGCTGGGTGTGCTGTGGAGAGTGTCTTCTCTG 120
QY 264 CTCCTCCCTCATGGCTGGCTGTGCTGTGGAGTGTCTTCTTATCGGAATGTGCGGGTTCAA 323
Db 121 CTCCTCCCTCATGGCTGGCTGTGCTGTGGAGTGTCTTCTTATCGGAATGTGCGGGTTCAA 180
QY 324 AAAGTCTTCAATGGCCATCTGAGATCAACAATGAGATCTTTTGTGATGCTATGAGAAC 383
Db 181 AAAGTCTTCAATGGCCATCTGAGATCAACAATGAGATCTTTTGTGATGCTATGAGAAC 240
QY 384 TCCACCTTCAAGAGTTTATGAGCTGGCCAGCCGAGGAGAGGAGCGCTGAAGTGTG 443
Db 241 TCCACCTTCAAGAGTTTATGAGCTGGCCAGCCGAGGAGAGGAGCGCTGAAGTGTG 300
QY 444 TACAATGAAGTCCCTGCTGGGTCCCTACACAAGAAGTGGCTGTAACTGCTTCACT 503

Db 301 TACAATGAAGTCCCTGCTGGGTCCCTTACCAAGAAGTGGCTGTAACTGCTTCACTGAG- 359
QY 504 GAGGCACTGTTCATGGCTTACTACTGTGTGAGAGTTTCAAGATCCCCCACTGGCAGAA 563
Db 360 ----- 359
QY 564 GAGTTGATTCGGCCATGGCTGTGAGAGATTGTAACTTGCACCCGAGACGAGCA 623
Db 360 ----- 359
QY 624 CTGAATCTTCTGTCTAATCTGTGTGGCTTCCCATTTGACCCCAAGATGCTGAG 683
Db 360 ----- 401
QY 684 AGGACTCAGACAACAAGCTGCAGTTTGGCTGCATGCCATGTGACAGAGTACACGC 743
Db 402 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 461
QY 744 TTCACCTACCCCTGGCTTCCCAACAGTCCCTACCCGGCGATGCCGCTGGCAGTGGCTC 803
Db 462 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 521
QY 804 CTGGCGGGGAGACGCCAGCTGTGTGAGACCTCACCTTCGAAGCTTGTATGTGCTGCC 583
Db 522 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 581
QY 864 TGTGATGAGCATGAGCAGTACCTGTACCGTGTATGATAGCTGAGCCCATGAAACC 923
Db 582 TGTGATGAGCATGAGCAGTACCTGTGTACCGTGTATGATAGCTGAGCCCATGAAACC 641
QY 924 CACGCTGTGTGGGTGTGTGTGACCTTCTACCTCTCTAACAACCTGACTTCTCTCC 983
Db 642 CACGCTGTGTGGGTGTGTGTGACCTTCTCTAACAACCTGACTTCTCTCC 701
QY 984 TCCCAAGAGCTCTTCTTGTGACAGCTGATACCAATGACCGGAGCATCTGCTT 1043
Db 702 TCCCAAGAGCTCTTCTTGTGACAGCTGATACCAATGACCGGAGCATCTGCTT 761
QY 1044 GAGGCCACTTCTTCCAGCTGCCCAAGATGAGAGCTGTGGCGCTTTTGTAGTGACCC 1103
Db 762 GAGGCCACTTCTTCCAGCTGCCCAAGATGAGAGCTGTGGCGCTTTTGTAGTGACCC 821
QY 1104 CAAAGGACATTTAGCAGCCCTTACTATCTACAGGCACTACCCGCCCAATCACTGAC 1163
Db 822 CAAAGGACATTTAGCAGCCCTTACTATCTACAGGCACTACCCGCCCAATCACTGAC 881
QY 1164 TGGAAATTCAAAGTGTGCCCAACAACCGGAAGTGTGGCTTCAAACTCTTCTATCTG 1223
Db 882 TGGAAATTCAAAGTGTGCCCAACAACCGGAAGTGTGGCTTCAAACTCTTCTATCTG 941
QY 1224 GTGAGCCCAACGTACAGTGGGCTCCTGCAACAAGACTATGTGAGATCAACGGGAG 1283
Db 942 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1001
QY 1284 AAGTACTGCGGTGAGAGCTCCAGTTTGTGTGACGACGACACAGCAAGATTACAGTC 1343
Db 1002 NNGTACTGCGGTGAGAGCTCCAGTTTGTGTGACGACGACACAGCAAGATTACAGTC 1061
QY 1344 CACTTCAATCTGATCTGCTGTACAGGACACCGGGTCTCTAGTGTGATCTCTCTAC 1403
Db 1062 CACTTCAATCTGATCTGCTGTACAGGACACCGGGTCTCTAGTGTGATCTCTCTAC 1121
QY 1404 GACTTCAACGACCCGCTGCCAGGAGTGTCAATGTGCAAGACTGACGCTGATCCGAAG 1463
Db 1122 GACTTCAACGACCCGCTGCCAGGAGTGTCAATGTGCAAGACTGACGCTGATCCGAAG 1181
QY 1464 GAACCTGCGTGGCAGCGCTGGGACAGCTGCCGATTTATGTAGTACGCTTACGCGA 1523
Db 1182 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1241
QY 1524 TGCATGCGACCCAGCTTACGTCGCAAAACAGTTCTGCAAGCCCTCTTCTGGGTC 1583

[illegible]

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Komoto, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Harama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Maehara, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

MEDLINE 11076861

PUBMED 11076861

REFERENCE

AUTHORS Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length CDNAS

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 11076861

PUBMED 11076861

REFERENCE

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, I., Hirokane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, S., Kawal, J., Koike, Y., Kondo, S., Komoto, H., Koude, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Satou, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sanjo, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takai-Akihira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan [E-mail: genome-research@riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

```

source
1. .2689
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D630022P10"
/db_xref="MGI:2422473"
/db_xref="taxon:10090"
/clone="D630022P10"
/tissue_type="kidney"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
1. .2689
/note="suppression of tumorigenicity 14 (colon carcinoma) (MGI:MGI:1338881, GB|NM_011176, evidence: BLASTN, 100%, match=720)"

ORIGIN

Query Match      30.8%; Score 956.4; DB 11; Length 2689;
Best Local Similarity 76.7%; Pred. No. 3.5e-177;
Matches 1415; Conservative 0; Mismatches 1; Indels 428; Gaps 3;

QY      1 CATGTAGACGGCTGCGCCGAGAGGACCAAGCGCTGTGAGACCGGCATCGGACCGCCAAA 60
Db      110 CATGTAGACGGCTGCGCCGAGAGGACCAAGCGCTGTGAGACCGGCATCGGACCGCCAAA 169
QY      61 CCGTGGTGGCAATCGGCGCCGAAAGCGCGAGGGGCTCTGAGACCTTCGCGCGGAGC 120
Db      170 CCGTGGTGGCAATCGGCGCCGAAAGCGCGAGGGGCTCTGAGACCTTCGCGCGGAGC 229
QY      121 TCAAGTACAACTCCGCGCTAGAGAACATGATGGCTTTGAGAGGGGTGTGAGCTTCCTGC 180
Db      230 TCAAGTACAACTCCGCGCTAGAGAACATGATGGCTTTGAGAGGGGTGTGAGCTTCCTGC 289
QY      181 CTGCGAACATGCCAAGAAAGTGAGAACGAGCGCCCGAGCGCTGGGTGTGCTGTGAG 240
Db      290 CTGCGAACATGCCAAGAAAGTGAGAACGAGCGCCCGAGCGCTGGGTGTGCTGTGAG 349
QY      241 CAGTCTGTGACGCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db      350 CAGTCTGTGACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 409
QY      301 ATTATCGGAATGCGGGGTTCAAAAAGTTCATATGCGCCATGAGAGATCAAAATGAGA 360
Db      410 ATTATCGGAATGCGGGGTTCAAAAAGTTCATATGCGCCATGAGAGATCAAAATGAGA 469
QY      361 TCTTCTGTGATGCGTATGAGAACTCCACTCCACAGAGTTTATCAG-CCTGCGCAGCCAG 419
Db      470 TCTTCTGTGATGCGTATGAGAACTCCACTCCACAGAGTTTATCAGCCCTGCGCAGCCAG 529
QY      420 GTGAAGAGAGCGCTGAAAGCTGTGTAATGAAAGTCCCTGTCTGGGTCCCTACCAAG 479
Db      530 GTGAAGAGAGCGCTGAAAGCTGTGTAATGAAAGTCCCTGTCTGGGTCCCTACCAAG 589
QY      480 AAGTCGGCTGTAACTGCGCTTCAAGTGAAGGCGAGTGCATCGCTTAATTACTGTGCAAGTTC 539
Db      590 AAGTCGGCTGTAACTGCGCTTCAAGTGAAGGCGAGTGCATCGCTTAATTACTGTGCAAGTTC 649
QY      540 AGCATCCCCCAACAAGTGGAGAGAGGTTGATCGCGCATGAGCTGTGGAAGGAGTTGTA 599
Db      650 AGCATCCCCCAACAAGTGGAGAGAGGTTGATCGCGCATGAGCTGTGGAAGGAGTTGTA 709
QY      600 ACATTGCGCAACCCCGAGACGCGGCACTGAATCTCTGTGCTAACATCTGTGTGCGCTTC 659
Db      710 ACATTGCGCAACCCCGAGACGCGGCACTGAATCTCTGTGCTAACATCTGTGTGCGCTTC 769
QY      660 ----- 659
Db      770 CGTAGTGGCGGGGCAAAACCAAGGTGGGAATGTGTGGCCAGACCTGTAAAGGGCGGGGCT 829
QY      660 ----- 659
Db      830 GGAAGAGCCACACTACAGGCCCAAGGTGTCTGTGAAGAGACTGTGCTGTGAGTGGCACT 889

```

```

QY      660 ----- 659
Db      890 CCTTTTGGTAATAGGAGGCAAGATTAAGAATATGTCTGGTAAATTTTGCTT 949
QY      660 -----CCATTGAACCCAGAAATGCTGACAGAGACTTACAGACA----- 696
Db      950 TCATGTTCACACCCATTACCCAGAAATGTGACAGAGACTACAGATATCAT 1009
QY      697 ----- 696
Db      1010 GGTGACCTGTGTGGAAGAAAAAGTGTCTTGCCCTTAAGCCATGAGGCTCCCTAAAGT 1069
QY      697 ----- 696
Db      1070 CAGATGTACTGATGAGCAAGATCTCTTTTCCCTTCGCGTCCAAAGCTCTTTC 1129
QY      697 ----- 696
Db      1130 TTGTGTCTACTGATTCGAGTGTGATCTGAGACGTAATGGGGGTAGAGCTAGGCGGGGG 1189
QY      697 -----ACAGTGCAGTTTTC 712
Db      1190 CGGGGGCTGCAAGGTTTACAGCTTCTTCCATCTCTCCACAGACAGCTGAGTTTTC 1249
QY      713 CCGTGAATGCCATGTGTGACAGATGACAGCTTCACTACCTGCTTCCCAACAGTCC 772
Db      1250 CCGTGAATGCCATGTGTGACAGATGACAGCTTCACTACCTGCTTCCCAACAGTCC 1309
QY      773 CTACCCGGCGCATGCCCCGCTGCAATGGGTCTGCGGGGGAGCGCCGACTGTGTGAG 832
Db      1310 CTACCCGGCGCATGCCCCGCTGCAATGGGTCTGCGGGGGAGCGCCGACTGTGTGAG 1369
QY      833 CCTCACCTTCGGAAGCTTGAATGTGCTCCCTGATATAGCATGACAGAGACCTGCTAC 892
Db      1370 CCTCACCTTCGGAAGCTTGAATGTGCTCCCTGATATAGCATGACAGAGACCTGCTAC 1429
QY      893 CGTGTATATAGCTCTGAGCCCGCATGGAACCCGACGCTGTGTGCGGCTGTGAGCACTT 952
Db      1430 CGTGTATATAGCTCTGAGCCCGCATGGAACCCGACGCTGTGTGCGGCTGTGAGCACTT 1489
QY      953 CTCACCTTCCTCAACCTGACTTCTCTCTCTCCAGAAAGCTCTTCTGTGACGCTGAT 1012
Db      1490 CTCACCTTCCTCAACCTGACTTCTCTCTCTCCAGAAAGCTCTTCTGTGACGCTGAT 1549
QY      1013 AACCAATCTGACCGGGGCAATCCGCTTTGAGGCGCACTTCTTCCAGAGTCCCAAGAT 1072
Db      1550 AACCAATCTGACCGGGGCAATCCGCTTTGAGGCGCACTTCTTCCAGAGTCCCAAGAT 1609
QY      1073 GAGCAGCTGTGCGGCTTTTGTAGTGAACCCCAAGGACATTTAGCAGCCCTATATCC 1132
Db      1610 GAGCAGCTGTGCGGCTTTTGTAGTGAACCCCAAGGACATTTAGCAGCCCTATATCC 1669
QY      1133 AGGCCTACTACCCGCCCAATCAACTGCAATGGAATATCAAGTGTGCCCAACCGGAA 1192
Db      1670 AGGCCTACTACCCGCCCAATCAACTGCAATGGAATATCAAGTGTGCCCAACCGGAA 1729
QY      1193 CGTGAAGGTGCGCTTAAACTCTTCTATCTGTGAGACCCCAACGTAACAGTGGGCTCTCG 1252
Db      1730 CGTGAAGGTGCGCTTAAACTCTTCTATCTGTGAGACCCCAACGTAACAGTGGGCTCTCG 1789
QY      1253 CACCAAGCATATGTGAGATCAACGAGGAGAACTACTGCGGTGAGAGTCCAGATTGT 1312
Db      1790 CACCAAGCATATGTGAGATCAACGAGGAGAACTACTGCGGTGAGAGTCCAGATTGT 1849
QY      1313 GGTGACAGCAACAGCAGCAAGATTAAGTCACTTCCATTCTGATCACTGTACAGGA 1372
Db      1850 GGTGACAGCAACAGCAGCAAGATTAAGTCACTTCCATTCTGATCACTGTACAGGA 1909
QY      1373 CACCGGTTCTAGCTAGTACTCTCTCTACAGCTCCAAAGAC 1416
Db      1910 CACCGGTTCTAGCTAGTACTCTCTCTACAGCTCCAAAGAC 1953

```

```

RESULT 7
BQ885551          944 bp      mRNA      linear      EST 16-ANG-2002
LOCUS             BQ885551
DEFINITION       AGENCOURT 8748027 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335237
5', mRNA sequence.
ACCESSION        BQ885551
VERSION          BQ885551.1 GI:22277569
KEYWORDS         EST.
SOURCE           Mus musculus (house mouse)
ORGANISM         Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 944)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
cDNA Library Preparation: Resgen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM13797 row: f column: 06
High quality sequence stop: 623.
Location/Qualifiers
1..944
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6335237"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 130"
/note="Organ: olocyts; Vector: pCMV-SPORT6.1; Site: 1;
EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size 1.95 kb. Constructed by
Resgen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      27.2%; Score 843.8; DB 13; Length 944;
Best Local Similarity 96.3%; Pred. No. 4.8e-155;
Matches 885; Conservative 0; Mismatches 32; Indels 2; Gaps 2;

QY 30 GCGTCTGAGACCGCGCATCGGACCCGCAAAACCATGGGTAGCAATGCGGCGCGCAAGGCC 89
Db 1 GCGTCTGAGACCGGCGATCGGACCGCAAAACCATGGGTAGCAATCGGCGCGCAAGGCC 60

QY 90 GGAAGGGGCTCTCAGACTTCGGGCGGGGACTCAAGTACAACTCCCGGCTAGAGAAATG 149
Db 61 GGAAGGGGCTCTCAGACTTCGGGCGGGGACTCAAGTACAACTCCCGGCTAGAGAAATG 120

QY 150 AATGCGCTTTGAGAGGGGTGAGAGTCTGCGCTCGCAACATGCGCAAGAAAGTGAAG 209
Db 121 AATGCGCTTTGAGAGGGGTGAGAGTCTGCGCTCGCAACATGCGCAAGAAAGTGAAG 180

QY 210 CGAGGCCCGAGGCGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 269
Db 181 CGAGGCCCGAGGCGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240

QY 270 CTATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 329
Db 241 CTATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 300

QY 330 TTCAATGGCCATCGAGATCACAATAGATCTTTCTGATGGGTATGAGAACTCCACC 389
Db 301 TTCAATGGCCATCGAGATCACAATAGATCTTTCTGATGGGTATGAGAACTCCACC 360

QY 390 TCACAGAGTTTATCAGCTCGGCGCAGCGAGGTGAAGAGAGGCGCTGAAGCTCTGTCAAT 449
Db 361 TCACAGAGTTTATCAGCTCGGCGCAGCGAGGTGAAGAGAGGCGCTGTGTCAAT 420

```

```

QY 450 GAAATCCCTGTCTCGAGTCCCTACCAAGAAGTGGCTGTAACTGCTTCACTAGTAGGC 509
Db 421 GAAATCCCTGTCTCGAGTCCCTACCAAGAAGTGGCTGTAACTGCTTCACTAGTAGGC 480

QY 510 AGTGTATGCTACTACTACTGTGTGAGATTTCAGATCCCTCCACACCTGGCAAGAGGTT 569
Db 481 AGTGTATGCTACTACTACTGTGTGAGATTTCAGATCCCTCCACACCTGGCAAGAGGTT 540

QY 570 GATCGGCGCATGCTGTGAGAGGAGTTGATGACATGGCCACCCCGAGCAGGCACTGA 629
Db 541 GATCGGCGCATGCTGTGAGAGGAGTTGATGACATGGCCACCCCGAGCAGGCACTGA 600

QY 630 TCTTGTGCTAAACATCTGTGTGAGCTTCCCATTTAGCCCGAGATGTGAGAGACT 689
Db 601 TCTTGTGCTAAACATCTGTGTGAGCTTCCCATTTAGCCCGAGATGTGAGAGACT 660

QY 690 CAGACACAAGCTGCAAGTTTGGCTCTGCAATGCCATGCTGACAGTACAGCTTCACT 749
Db 661 CAGACACAAGCTGCAAGTTTGGCTCTGCAATGCCATGCTGACAGTACAGCTTCACT 720

QY 750 ACCCTGGCTTCCCGCAAGTCCCTACCCCGGCGCATGCCGCTGCAAGTGGTCTGCGG 809
Db 721 ACCCTGGCTTCCCGCAAGTCCCTACCCCGGCGCATGCCGCTGCAAGTGGTCTGCGG 780

QY 810 GGGAGCGCGACTCTGTGTGAGCTTACCTTCCAGAGCTTTGATGTGCTCTGTGAT 869
Db 781 GGGAGCGCGACTCTGTGTGAGCTTACCTTCCAGAGCTTTGATGTGCTCTGTGAT 840

QY 870 GAGCAT-GGCATGACCTGTGACCGT-GTATATAGCCTGACGCCCATGGAACCCACG 927
Db 841 GAGCATGAGGATGATGCTTGGCCCGCGGTGTATGATACCTGAGCCCATGGAACCCCA 900

QY 928 CTGTGTGCGGCGCTGTGTGG 946
Db 901 CCTGTGTGTGCGGCGCTTG 919

RESULT 8
B1688367          884 bp      mRNA      linear      EST 18-SEP-2001
LOCUS             B1688367
DEFINITION       603316592F1 NCI_CGAP_Mame6 Mus musculus cDNA clone IMAGE:5356487 5',
mRNA sequence.
ACCESSION        B1688367
VERSION          B1688367.1 GI:15650996
KEYWORDS         EST.
SOURCE           Mus musculus (house mouse)
ORGANISM         Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 884)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1906 row: h column: 24
High quality sequence stop: 831.
Location/Qualifiers
1..884
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5356487"
/sex="female, virgin"

```

```

/tissue type="infiltrating ductal carcinoma"
/dev stage="5 months"
/lab host="DH10B"
/clone_id="NCT_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

```

```

Query Match      26.8%; Score 833.6; DB 12; Length 884;
Best Local Similarity 98.6%; Pred. No. 4.8e-153;
Matches 872; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

```

```

QY 1165 GGAATATCAAGTGTCCCAACACCGGAAGCTGAAGTGCCTTCAACTCTTATCTG 1224
DB 1 GGAATATCAAGTGTCCCAACACCGGAAGCTGAAGTGCCTTCAACTCTTATCTG 60
QY 1225 TGGACCCCAACGATGAGTGGCTCCTGACCAAGGACTATGTGAGATCAACGGGAGA 1284
DB 61 TGGACCCCAACGATGAGTGGCTCCTGACCAAGGACTATGTGAGATCAACGGGAGA 120
QY 1285 AGTACTGCGGTGAGAGTCCCAAGTTGTGTGAGAGCAACAGCAGATTACAGTCC 1344
DB 121 AGTACTGCGGTGAGAGTCCCAAGTTGTGTGAGAGCAACAGCAGATTACAGTCC 180
QY 1345 ACTTCATTCTATCTACTGTACAGGACACCGGTTCTTACTGATGATCTCTCTAGC 1404
DB 181 ACTTCATTCTATCTACTGTACAGGACACCGGTTCTTACTGATGATCTCTCTAGC 240
QY 1405 ACTCCAGACCCGTCGCCAGGATGTATGTGACAGAGTGCAGCTGATCCGAAG 1464
DB 241 ACTCCAGACCCGTCGCCAGGATGTATGTGACAGAGTGCAGCTGATCCGAAG 300
QY 1465 AACTGCGTGCAGACGCGTGGGACAGCTGCCGATTAATGATGAGCTTACTGCCAT 1524
DB 301 AACTGCGTGCAGACGCGTGGGACAGCTGCCGATTAATGATGAGCTTACTGCCAT 360
QY 1525 GGAATGCCACCAACCAAGTTCAGTGTCAAAACCAAGTTCGCAAGCCCTCTTCTGGTCT 1584
DB 361 GGAATGCCACCAACCAAGTTCAGTGTCAAAACCAAGTTCGCAAGCCCTCTTCTGGTCT 420
QY 1585 GTGACAGTGTCAACGATGTGGGAGACGAGTGAAGAGGAGGCTGCAGCTGTCTGCTG 1644
DB 421 GTGACAGTGTCAACGATGTGGGAGACGAGTGAAGAGGAGGCTGCAGCTGTCTGCTG 480
QY 1645 GGAATTTCAAGTGTTCATATGGGAAAGTGTCTCCCTCAGAGCAGAAAGTGAATGGAGAG 1704
DB 481 GGAATTTCAAGTGTTCATATGGGAAAGTGTCTCCCTCAGAGCAGAAAGTGAATGGAGAG 540
QY 1705 ACAACTGTGAGATGGGTCTGACGAGAGCTTCATGTGACAGCGTGAATGTCTCTTGA 1764
DB 541 ACAACTGTGAGATGGGTCTGACGAGAGCTTCATGTGACAGCGTGAATGTCTCTTGA 600
QY 1765 CCAAAATATCTACCGCTGCCAAATATGGCTCTGTCTGAG-CAAGGGGAAACCTTAGATGT 1823
DB 601 CCAAAATATCTACCGCTGCCAAATATGGCTCTGTCTGAGCAAGGGGAAACCTTAGATGT 660
QY 1824 GATGGGAAGACGAGCTGTAGCAGTGGCTCCGATGAGAAAACTGTGACTGTGGCTGCA 1883
DB 661 GATGGGAAGACGAGCTGTAGCAGTGGCTCCGATGAGAAAACTGTGACTGTGGCTGCA 720
QY 1884 TCCCTTACCAACAGAGCTGCGG- TGGTGTGTGACGAAATGCGAGCGAGGAGTGGCC 1942
DB 721 TCCCTTACCAACAGAGCTGCGG- TGGTGTGTGACGAAATGCGAGCGAGGAGTGGCC 780
QY 1943 CTGGCAGGTGAGCTCCAGCGCTGGGCGCAGGGCCACTTGTGTGGGCTGCG- TCAATCT 2001
DB 781 CTGGCAGGTGAGCTCCAGCGCTGGGCGCAGGGCCACTTGTGTGGGCTGCGCTTCAATCT 840
QY 2002 CTCTGACTGTGTCTCTGTGACAGCTCAATGCTTTCAAGATGAC 2045
DB 841 CTCTGACTGTGTCTCTGTGACAGCTCAATGCTTTCAAGATGAC 884

```

```

RESULT 9
BI690615 803 bp mRNA linear EST 18-SEP-2001
LOCUS 603314210F1 NCT_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5354067 5'
DEFINITION
mRNA sequence.
BI690615
ACCESSION BI690615.1 GI:15653244
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

```

```

REFERENCE
AUTHORS
TITLES
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Inocyte Genomics, Inc.
CDNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI at:
http://image.llnl.gov
Plate: LLAM11900 row: d column: 04
High quality sequence stop: 755.

```

FEATURES

source

```

1..803
Location/Qualifiers

```

```

/organism="Mus musculus"
/mol type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5354067"
/sex="female, virgin"
/tissue type="infiltrating ductal carcinoma"
/dev stage="5 months"
/lab host="DH10B"
/clone_id="NCT_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

```

ORIGIN

```

Query Match      25.4%; Score 788.8; DB 12; Length 803;
Best Local Similarity 99.6%; Pred. No. 3e-144;
Matches 801; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

```

QY 654 GCCCTCCCATTTGACCCCAAGATGCTGACAGAGGACTCAGAGCAACAGCTGAGTTGGC 713
DB 1 GCCCTCCCATTTGACCCCAAGATGCTGACAGAGGACTCAGAGCAACAGCTGAGTTGGC 60
QY 714 CTGCATGCCATGTGTGACAGAGTGAACAGCTTCACTAACCCTGCTCCCAACAGTCC 773
DB 61 CTGCATGCCATGTGTGACAGAGTGAACAGCTTCACTAACCCTGCTCCCAACAGTCC 120
QY 774 TACCCGGGCAATGCCCGCTGCCAGTGGTCTGCGGGGGAGACGCCGACTCTGTGTGAGC 833
DB 121 TACCCGGGCAATGCCCGCTGCCAGTGGTCTGCGGGGGAGACGCCGACTCTGTGTGAGC 180
QY 834 CTCACCTTCGGAAGCTTGTGATGTGCTCCCTGTATGAGCATGTGAGAGTACCTGTGACC 893
DB 181 CTCACCTTCGGAAGCTTGTGATGTGCTCCCTGTATGAGCATGTGAGAGTACCTGTGACC 240
QY 894 GTGTATGATAGCTGAGCCCATGGAACCCACAGCTGTGTGTGGCTGTGTGACCTTC 953
DB 241 GTGTATGATAGCTGAGCCCATGGAACCCACAGCTGTGTGTGGCTGTGTGACCTTC 300
QY 954 TCACCTTCTTACACCTGACTTCTCTCTCCAGAAAGTCTTCTTGTACGCTGATA 1013

```


Db 956 CAGGGGCCCCGTGGAGGGGAGCACTCCCGGCGTGG 990

RESULT 11
LOCUS BG919617
DEFINITION 602823173F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4951962 5',
mRNA sequence.
ACCESSION BG919617
VERSION BG919617
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 762)
NIH-MGC <http://imgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: LLAM10909 row: a column: 19
High quality sequence stop: 727.
location/Qualifiers
1..762
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4951962"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies, Investigator
providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Query Match 24.5%; Score 760.4; DB 12; Length 762;
Best Local Similarity 99.9%; Pred. No. 1.1e-138;
Matches 761; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 682 AGAGGACTCAGAGCAACAGTGTGCTTGGCCGTCAGTGGTGCACAGTGCAC 741
Db 1 AGAGGACTCAGAGCAACAGTGTGCTTGGCCGTCAGTGGTGCACAGTGCAC 60
QY 742 GCTTCACTACCCCTGGCTTCCCAACAGTCCCTACCCGCGCATGCCGCTGCCAGTGGG 801
Db 61 GCTTCACTACCCCTGGCTTCCCAACAGTCCCTACCCGCGCATGCCGCTGCCAGTGGG 120
QY 802 TCCGCGGGGGGAGCCGACCTCTGTGCTGAGCTCACCCTTCGAGCTTTGATGTCGCTC 861
Db 121 TCCGCGGGGGGAGCCGACCTCTGTGCTGAGCTCACCCTTCGAGCTTTGATGTCGCTC 180
QY 862 CTTGTATGAGCATGGAGTGAAGTCCGATCCGATGATAGCTAGAGCCCATGGAAC 921
Db 181 CTTGTATGAGCATGGAGTGAAGTCCGATCCGATGATAGCTAGAGCCCATGGAAC 240
QY 922 CCAAGCTGTGTGGGCTGTGTGGACCTTCTCACCCTCTCAACACTGACTTTCTCT 981
Db 241 CCAAGCTGTGTGGGCTGTGTGGACCTTCTCACCCTCTCAACACTGACTTTCTCT 300

QY 982 CTTCCAGAACGCTCTCTGTGTCAGCTGTATACCAATCTGACGGGCGCATCTGGCT 1041
Db 301 CTTCCAGAACGCTCTCTGTGTCAGCTGTATACCAATCTGACGGGCGCATCTGGCT 360
QY 1042 TTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCTGTGGGCTTTTGTAGTACA 1101
Db 361 TTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCTGTGGGCTTTTGTAGTACA 420
QY 1102 CCCAAGGACATTTAGCAGCCCTTACTATTCAGGCCACTACCCGCCCAACATCACTGCA 1161
Db 421 CCCAAGGACATTTAGCAGCCCTTACTATTCAGGCCACTACCCGCCCAACATCACTGCA 480
QY 1162 CATGAATATCAAGGAGCCCAACACCGGAAGGTGAAGGTGGCTTCAACTTTCTATC 1221
Db 481 CATGAATATCAAGGAGCCCAACACCGGAAGGTGAAGGTGGCTTCAACTTTCTATC 540
QY 1222 TGGTGACCCCAACGATCAGTGGGCTCTGACCAAGACATATGTGAGATCAACGGGG 1281
Db 541 TGGTGACCCCAACGATCAGTGGGCTCTGACCAAGACATATGTGAGATCAACGGGG 600
QY 1282 AAGAGTACTGCGGTGAGAGTCCAGTTTGTGTGAGCAGCAACAGCAAGATTACG 1341
Db 601 AAGAGTACTGCGGTGAGAGTCCAGTTTGTGTGAGCAGCAACAGCAAGATTACG 660
QY 1342 TCCACTTCATCTGATCACTGTCACGGAACCGGGTCTTAGTGATACCTCTCT 1401
Db 661 TCCACTTCATCTGATCACTGTCACGGAACCGGGTCTTAGTGATAGTACTCTCTCT 720
QY 1402 ACGACTCCAAAGACCCGTGCCAGGAGTGTTCATGTGCAGAGA 1443
Db 721 ACGACTCCAAAGACCCGTGCCAGGAGTGTTCATGTGCAGAGA 762

RESULT 12
LOCUS CB208872
DEFINITION 874 bp mRNA linear EST 16-May-2003
IMAGE:30241011 5', mRNA sequence.
ACCESSION CB208872
KEYWORDS CB208872.2 GI:29133712
EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 874)
NIH-MGC <http://imgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT On Feb 4, 2003 this sequence version replaced gi:28250435.
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. David Rowe and Dr. Mina
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM0308 row: g column: 04
High quality sequence stop: 655.
location/Qualifiers
1..874
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30241011"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 164"
/note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Non-normalized full-length enriched library from pooled
mouse embryonic limb, maxilla and mandible, day 10.5 and
11.5 (size selected for the 0.5-1 kb fragments) Cloned

FEATURES
source

ORIGIN

directionally, priming method: Oligo-dT. cDNA enrichment:
 21k bp, Average insert size 1.8k bp, Priming sequence:
 5'-GACTAGTCTAGATCGAGGCGGCGCCCTT-3', Tissue contributed
 by, David Rowe, Library constructed by Resgen, Invitrogen
 Corp."

Query Match	24.4%	Score 759.4	DB 14	Length 874	
Best Local Similarity	98.9%	Pred. No. 1.7e-138			
Matches 785	Conservative 0	Mismatches 7	Indels 2	Gaps 2	
QY	1387	CTGAGTACCTCTCTACGACATCCAAACGACCCGTCAGGATGTTTCATGTGCAAGACTG	1446		
DB	1	CTGAGTACCTCTCTCTACGACATCCAAACGACCCGTCAGGATGTTTCATGTGCAAGACTG	60		
QY	1447	GACGGTGCATCCGAAAGAACTGCGCTGCGACGCGCTGGGCACTCCCGGATTAATAGG	1506		
DB	61	GACGGTGCATCCGAAAGAACTGCGCTGCGACGCGCTGGGCACTCCCGGATTAATAGG	120		
QY	1507	ATGAGGGTATCGGCGATGGAATGCAACCCACGATTCAGTGCATGATGCAAGTCTGCA	1566		
DB	121	ATGAGGGTATCGGCGATGGAATGCAACCCACGATTCAGTGCATGATGCAAGTCTGCA	180		
QY	1567	AGCCCTCTCTCTGCTGCTGCAAGTGTCAACGACTGTGGGACGGAAGTACGAGGAGG	1626		
DB	181	AGCCCTCTCTCTGCTGCTGCAAGTGTCAACGACTGTGGGACGGAAGTACGAGGAGG	240		
QY	1627	GCTGACGCTGCTCTGCTGCTGCAAGTGTTCATGTTCCATGAGGAAAGTGTCTCCCTCAAGCC	1686		
DB	241	GCTGACGCTGCTCTGCTGCTGCAAGTGTTCATGTTCCATGAGGAAAGTGTCTCCCTCAAGCC	300		
QY	1687	AGAAAGTATATGGAAGGAAAGCACTGTGGAGATGGGTCTGACGAGGCTTCATGTGACAGC	1746		
DB	301	AGAAAGTATATGGAAGGAAAGCACTGTGGAGATGGGTCTGACGAGGCTTCATGTGACAGC	360		
QY	1747	TGAATGCTGCTCTGCTGCTGCAACCAATATACCTACGCTGCCAAATGGCTGTGTGAGCA	1806		
DB	361	TGAATGCTGCTCTGCTGCTGCAACCAATATACCTACGCTGCCAAATGGCTGTGTGAGCA	420		
QY	1807	AGGGCAACCTCTGATGTGATGGGAAAGCGACTGTGACGATGGCTCCGATGAGAAAACT	1866		
DB	421	AGGGCAACCTCTGATGTGATGGGAAAGCGACTGTGACGATGGCTCCGATGAGAAAACT	480		
QY	1867	GTAATGCTGCTGCTGCTGCTGCAACCAAGGCTCCGCTGCTGTGTGAGCAAGATGCG	1926		
DB	481	GTAATGCTGCTGCTGCTGCTGCAACCAAGGCTCCGCTGCTGTGTGAGCAAGATGCG	540		
QY	1927	ACGAGGGCGAGTGGCCCTGGGAGGTGAGCTCCACGCTGGGCGAGGCGCACTTGTGTG	1986		
DB	541	ACGAGGGCGAGTGGCCCTGGGAGGTGAGCTCCACGCTGGGCGAGGCGCACTTGTGTG	600		
QY	1987	GGGCTGCTGCTATCTCTCTGACGCTGCTGTCTGACGCTCATTTGCTTCAGGATGCA	2046		
DB	601	GGGCTGCTGCTATCTCTCTGACGCTGCTGTCTGACGCTCATTTGCTTCAGGATGCA	660		
QY	2047	AAATTTCAATGACTGACATACAGATGATGAGCGCTTCTGCTGCTGCTGCAAGCA	2106		
DB	661	AAATTTCAATGACTGACATACAGATGATGAGCGCTTCTGCTGCTGCTGCAAGCA	719		
QY	2107	GCAAGCGAGTGGCTCTGGGGGTGAGAGAGTGAAGCTCAAAAGTATCAT -CACCACTT	2165		
DB	720	GCAAGCGAGTGGCTCTGGGGGTGAGAGAGTGAAGCTCAAAAGTATCAT -CACCACTT	779		
QY	2166	TCCTTCAATGATTT 2179			
DB	780	TCCTTCAATGATTT 793			

RESULT 13
 B1664081 798 bp mRNA linear EST 12-SEP-2001
 LOCUS B1664081 603283835P1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5323485 5',
 DEFINITION mRNA sequence.

ACCESSION

B1664081
 B1664081.1 GI:15578314
 EST.
 Mus musculus (house mouse)
 SOURCE
 Mus musculus

REFERENCE

1 (bases 1 to 798)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1M11820 row: 1 column: 22
 High quality sequence stop: 771.
 Location/Qualifiers

FEATURES

1..798
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5323485"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam6"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Query Match	24.1%	Score 747.2	DB 12	Length 798	
Best Local Similarity	98.6%	Pred. No. 4.3e-136			
Matches 785	Conservative 0	Mismatches 8	Indels 3	Gaps 3	
QY	16	CCCGAGGAGCAACGCTCTGAGACCGGCGATGGAATCCGCCAAACATGCGTGGCATC	75		
DB	1	CCCGAGGAGCAACGCTCTGAGACCGGCGATGGAATCCGCCAAACATGCGTGGCATC	60		
QY	76	GGGGCCGCAAGGCGGAGGGGGCTCTCAGGACTTCGGCGGGGACTCAAGTCAACTCC	135		
DB	61	GGGGCCGCAAGGCGGAGGGGGCTCTCAGGACTTCGGCGGGGACTCAAGTCAACTCC	120		
QY	136	GACTAGAGAACATGATGAGCTTTGAGAGAGGTGTGAGATTCTGCTGCGCAACATGCCA	195		
DB	121	GACTAGAGAACATGATGAGCTTTGAGAGAGGTGTGAGATTCTGCTGCGCAACATGCCA	180		
QY	196	AGAAAGTGAAGAGCGAGCCCGAGGGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	255		
DB	181	AGAAAGTGAAGAGCGAGCCCGAGGGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	240		
QY	256	TCCTCTGCTCTGCTCTGATGAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	315		
DB	241	TCCTCTGCTCTGCTCTGATGAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	300		
QY	316	GGTTCAAAAAGTCTTCAATGAGCCATCTGAGGATCACAATGAGATCTTCTGATGGCT	375		
DB	301	GGTTCAAAAAGTCTTCAATGAGCCATCTGAGGATCACAATGAGATCTTCTGATGGCT	360		
QY	376	ATGAGAACTCACTCCACAGAGTTTATCAGCTTGGCCAGCCAGGTGAGAGAGGCGCTGA	435		
DB	361	ATGAGAACTCACTCCACAGAGTTTATCAGCTTGGCCAGCCAGGTGAGAGAGGCGCTGA	420		
QY	436	AGCTGCTGTACATGAAATCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	495		

Db 421 AGCTGCTGATCAATGAAGTCCCTGTCTCTGAGTCCCTACCAAGAAATGAGTACTG 480

QY 496 CTTTCACTGAGGAGGAGTGTATCGCTACTACTACTAGTGTAGAGTTCAGAGTCCCCCAACC 555

Db 481 CTTTCACTGAGGAGGAGTGTATCGCTACTACTACTAGTGTAGAGTTCAGAGTCCCCCAACC 540

QY 556 TGGCAAGAAAGTTGATTCGGCCATGAGTGTGAGAGGAGTTGTAACATTGCCACCCGAG 615

Db 541 TGGCAAGAAAGTTGATTCGGCCATGAGTGTGAGAGGAGTTGTAACATTGCCACCCGAG 600

QY 616 CACGGGACCTGAATCTCTGTCTGTAACTCTGTGTGGCTTCCCATTTAGACCCAGAA 675

Db 601 CACGGGACCTGAATCTCTGTCTGTAACTCTGTGTGGCTTCCCATTTAGACCCAGAA 660

QY 676 TGTGTGAGAGGAGTCTAGAGCAACAGCTGAGTTTGGCTTCCCTGATG-CCCATGGTGGAGA 734

Db 661 TGTGTGAGAGGAGTCTAGAGCAACAGCTGAGTTTGGCTTCCCTGATG-CCCATGGTGGAGA 720

QY 735 GTG-AACGCTTCTACTACCTCTGCTTCCCAACAGTCTCCCTAACCCGCGCATGCCGCTG 793

Db 721 GTGAAACGCTTCTACTACCTCTGCTTCCCAACAGTCTCCCTAACCCGCGCATGCCGCTG 779

QY 794 CCAAGTGGTCTGCGG 809

Db 780 GCAAGTGGTCTGCGG 795

RESULT 14
BY756896
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BY756896 746 bp mRNA linear EST 17-DEC-2002
BY756896 RIKEN full-length enriched, blastocyst Mus musculus cDNA
Clone 11C0035F06 3', mRNA sequence.
BY756896
BY756896.1 GI:27190109
EST.

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogamathi; Muridae; Murinae; Mus.
1 (bases 1 to 746)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, A.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bull, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltara, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Nunata, K., Okido, T., Pavani, W.J., Petrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shmadel, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wysshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, I., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Landet, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED

Nature 420, 563-573 (2002)
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
12466951

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, I., Kojima, Y.,
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Nunata, K., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watanabe, A., Watanabe, M., and Hayashizaki, Y.

Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichipillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source

Location/Qualifiers
1..746
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="11C0035F06"
/dev_stage="blastocyst"
/clone_lib="RIKEN full-length enriched, blastocyst"

ORIGIN

Query Match 23.8%; Score 739.2; DB 13; Length 746;
Best Local Similarity 99.3%; Pred. No. 1.6e-134;
Matches 741; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2343 CTGATCTGCAAGAGGTGAGATCCGTGTCATCAACGACCACTGTGAGAGCACTTCATG 2402

Db 1 CTGATCTGCAAGAGGTGAGATCCGTGTCATCAACGACCACTGTGAGAGCACTTCATG 60

QY 2403 CCGAGAGATCACCCACGATGATGTGTGTGCTTCTCACTGAGGAGTGTGACTCC 2462

Db 61 CCGAGAGATCACCCACGATGATGTGTGTGCTTCTCACTGAGGAGTGTGACTCC 120

QY 2463 TGGCAGGAGTCTGTGTGCGCCCTTGTCAAGCCGGGAGAAAGATGGCGAATGTTCCAG 2522

Db 121 TGGCAGGAGTCTGTGTGCGCCCTTGTCAAGCCGGGAGAAAGATGGCGAATGTTCCAG 180

QY 2523 GCTGTGTGAGAGTGGAGTGAAGAGGCTGCTCAGAGAAACAGCCGAGGTGTACCA 2582

Db 181 GCTGTGTGAGAGTGGAGTGAAGAGGCTGCTCAGAGAAACAGCCGAGGTGTACCA 240

QY 2583 AGGTCCTGTAGTTGGGAGTGTATCAAGAGCACTGGGGTATATACAGATGAGACG 2642

Db 241 AGGTCCTGTAGTTGGGAGTGTATCAAGAGCACTGGGGTATATACAGATGAGACG 300

QY 2643 ACAGCCGACCAACAAACCCACAGAGATGCCGACATGCAACCTGTATACAGAGAGA 2702

Db 301 ACAGCCGACCAACAAACCCACAGAGATGCCGACATGCAACCTGTATACAGAGAGA 360

QY 2703 AACTGACGACATTTATGCTGTGCTCTCCCCCCCCCAACAACCCAGACTGGAAGCTGC 2762
 DB 361 AACTGACGACATTTATGCTGTGCTCTCCCCCCCCCAACAACCCAGACTGGAAGCTGC 420
 QY 2763 ATCCCTAGACATCAGAGTCTTCTCAAGTGGAGCCCTCCTCAGAGTTGGAGAGAACTT 2822
 DB 421 ATCCCTAGACATCAGAGTCTTCTCAAGTGGAGCCCTCCTCAGAGTTGGAGAGAACTT 480
 QY 2823 GCGTGTAGAGGCGCCAGCTGCGGCGCAAGGTTTATGCGAGCTTCCCTCTAGCCCT 2882
 DB 481 GCGTGTAGAGGCGCCAGCTGCGGCGCAAGGTTTATGCGAGCTTCCCTCTAGCCCT 540
 QY 2883 GAGCTGGTGAAGATATGCTGTCCCGAGAGCTGCTCAACTGATTTAGCTCCCG 2942
 DB 541 GAGCTGGTGAAGATATGCTGTCCCGAGAGCTGCTCAACTGATTTAGCTCCCG 600
 QY 2943 GAGCCCTATGGAGAGGAGGCTCAGGCTCACTCTTTTCAAGAAAGCCAGCCCTAGGAAC 3002
 DB 601 GAGCCCTATGGAGAGGAGGCTCAGGCTCACTCTTTTCAAGAAAGCCAGCCCTAGGAAC 660
 QY 3003 CCAGAGAAAGAGTGTACTTAAAGCTGAATTTGTTTGTCTGTCAGAGGCTGGTATT 3062
 DB 661 CCAGAGAAAGAGTGTACTTAAAGCTGAATTTGTTTGTCTGTCAGAGGCTGGTATT 720
 QY 3063 GAGAGTAAACATTTATTTCTTTT 3088
 DB 721 GAGAGTAAACATTTATTTCTTTT 746

RESULT 15

BI416218 846 bp mRNA linear EST 14-AUG-2001
 LOCUS 602989878F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5145762 5',
 DEFINITION mRNA sequence.

ACCESSION BI416218
 VERSION BI416218
 KEYWORDS GI:15177141
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 846)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: CGAPs-rc@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bernaldo, Ph.D.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11359 row: 1 column: 19
 High quality sequence start: 20
 High quality sequence stop: 793.
 Location/Qualifiers

FEATURES
 source

1..846
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech 1r"
 /db_xref="taxon:10090"
 /clone="IMAGE:5145762"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NCI CGAP_Lu33"
 /note="Organ: lung; Vector: pRTT3-Pac (Pharmacia) with a modified polylinker; Site 1: NCI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5'

ORIGIN

TGTACCAATCTGAAGTGGAGGCGCCCTCTGTTTTTTTTTTT 3'.
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bernaldo.

Query Match 23.8%; Score 739.2; DB 12; Length 846;
 Best Local Similarity 97.1%; Pred. 1.6e-134;
 Matches 816; Conservative 0; Mismatches 18; Indels 6; Gaps 6;

QY 1687 AGAAGTATATGGAGAGCAACTGTGAGATGGTGTGACGAGGCTTATGTGACAGG 1746
 DB 1 AGATCCGACAGAGTATAGGACACTGTGAGATGGTGTGACGAGGCTTATGTGACAGG 60
 QY 1747 TGAATGTCCTCTTTCACCAATATACCTAACCGCTGCCAAATGGGCTCTGTAGACA 1806
 DB 61 TGAATGTCCTCTTTCACCAATATACCTAACCGCTGCCAAATGGGCTCTGTAGACA 120
 QY 1807 AGGCAACCCCTGAGTGTATGGAGACGAGCTGTAGACATGGCTCCGATGAGAAAACT 1866
 DB 121 AGGCAACCCCTGAGTGTATGGAGACGAGCTGTAGACATGGCTCCGATGAGAAAACT 180
 QY 1867 GTGACTGTGGCTGCGATCTTTACCAACAGGCTGCGTGTGTGTGCGACGAATGCCG 1926
 DB 181 GTGACTGTGGCTGCGATCTTTACCAACAGGCTGCGTGTGTGTGCGACGAATGCCG 240
 QY 1927 AGGAGGGCGAGTGGCCCTGGCAGGTGAGCTCCACGCGCTGGGCGACAGGCACTTGATG 1986
 DB 241 AGGAGGGCGAGTGGCCCTGGCAGGTGAGCTCCACGCGCTGGGCGACAGGCACTTGATG 239
 QY 1987 GGGCTGTGCTCATCTCTCTGAGCTGCTGTGAGCTCATTTGCTTTCAGATGACA 2046
 DB 300 GGGCTGTGCTCATCTCTCTGAGCTGCTGTGAGCTCATTTGCTTTCAGATGACA 359
 QY 2047 AAAATTTCAGTACTAGATCAACATGAGACAGGCTTCCGGTCTGTGAGACAGA 2106
 DB 360 AAAATTTCAGTACTAGATCAACATGAGACAGGCTTCCGGTCTGTGAGACAGA 419
 QY 2107 GCAGAGGCACTGTGCTGGGGTGCAGAGGCTGAAGCTCAACGATATCAACCACTT 2166
 DB 420 GCAGAGGCACTGTGCTGGGGTGCAGAGGCTGAAGCTCAACGATATCAACCACTT 479
 QY 2167 CTTTCATGATTTACCTTGACTATGACATGCTTGTGAGCTGAGAGTGGAGTGG 2226
 DB 480 CTTTCATGATTTACCTTGACTATGACATGCTTGTGAGCTGAGAGTGGAGTGG 539
 QY 2227 AGTACAGACCTGTGTGCGGCCCATCTGCTGCTGATGTGTAACCATGCTTCCCTG 2286
 DB 540 AGTACAGACCTGTGTGCGGCCCATCTGCTGCTGATGTGTAACCATGCTTCCCTG 599
 QY 2287 GC-AAGGCATCTGGGTGCAGAGGCTGGGGGCACAAAGAGGAGGT-ACCGAGCGCT 2344
 DB 600 GCAGAGGCACTGTGCTGGGGTGCAGAGGCTGGGGGCACAAAGAGGAGGTCAACCGAGCGCT 659
 QY 2345 GATCTTCAGAGAGGAGTGTGATCTGTGATC-AAAGAGACCACTGTGAGAGCTCATGTC 2403
 DB 660 GATCTTCAGAGAGGAGTGTGATCTGTGATC-AAAGAGACCACTGTGAGAGCTCATGTC 719
 QY 2404 CGCAGCAGATCAACCCAGATATGATGTGTGTTTCTCAGTGGGGGTGTGACATCT 2463
 DB 720 CGCAG-AGATCAACCCAGATATGATGTGTGTTTCTCAGTGGGGGTGTGACATCT 778
 QY 2464 GCCAGAGTGACTGTGTGCGCCCTTGTCAAGCGGAGAAAGATGGCGAATGTTCCAGG 2523
 DB 779 G-CAGGAGTGTGTGTGCGCCCTTGTCAAGCGGAGAAAGATGTTCCAGG 837

Search completed: September 30, 2004, 16:31:28
 Job time : 7588 secs